

## Template feedback Hackathon#1

Notions évaluées : Extraction, traitement et consolidation de données

---

### Feedback – Hackathon

**Nom des élèves** : Ghizlane

**Nom du projet** : Attrition chez IBM

**Nom du professeur évaluateur** : Clara Martinez

**Date** : 03/07/2025

---

### Feedback global (2 phrases maximum)

#### Points positifs :

→ **Excellente structuration et démarche analytique** : Le plan du notebook est un modèle de workflow en data science (EDA → Prétraitement → Analyse → Visualisations → Recommandations). Il est logique, clair et suivi avec rigueur. Le code est très propre, bien segmenté et facile à lire.

→ **Une Analyse Exploratoire (EDA) très riche et pertinente** : C'est le grand point fort du projet. Chaque variable clé a été analysée méthodiquement. L'utilisation systématique de graphiques (**countplot**, **boxplot**) couplée à des tableaux de pourcentages (**crosstab**) pour quantifier les observations est une excellente pratique. Les interprétations sous chaque graphique sont justes et montrent une très bonne capacité à transformer les données en insights.

→ **Dépassement des attentes avec des analyses avancées** : Le projet ne s'est pas arrêté à une EDA classique. L'utilisation de **pivot\_table** pour croiser plusieurs variables (ex: **JobRole x Distance x Attrition**) est très pertinente. L'intégration d'un test statistique comme l'ANOVA pour vérifier des hypothèses et le clustering en bonus démontrent une réelle curiosité et une solide compréhension des outils.

#### Points à améliorer :

→ **Synthèse des insights** : L'analyse est très détaillée, avec une conclusion sous chaque graphique. Pour aller plus loin, une petite section "Synthèse de l'EDA" à la fin de la partie 2 aurait permis de regrouper les 3 ou 4 facteurs les plus importants identifiés (ex: "Notre analyse exploratoire pointe vers un profil à risque : les commerciaux juniors, célibataires, faisant des heures supplémentaires et insatisfaits de leur environnement."). Cela renforce l'impact du message.

→ **Peaufiner les visualisations** : Le travail de visualisation est globalement excellent, mais un petit détail a été noté. Sur le barplot montrant la corrélation des variables avec l'attrition (Partie 5), il manque une légende (colorbar) pour indiquer à quoi correspondent les

couleurs (le rouge pour les corrélations négatives, le bleu/chaud pour les positives). C'est un détail, mais important pour une lecture immédiate du graphique.

---

### Évaluation notée (notes sur 5)

*(Ces notes servent à départager les groupes pour la sélection du meilleur projet)*

Critère	Note /5	Commentaire
Clarté du code et commentaires pertinents	[ 5 ]	Code bien segmenté, clair, lisible et ordonné. L'utilisation de titres et de cellules Markdown est parfaite.
Utilisation efficace des notions vues pendant le bootcamp (Pandas, visualisation, EDA, etc.)	[ 4 ]	Excellente application des fondamentaux (Pandas, Seaborn) et utilisation de techniques plus avancées (pivot_table, clustering). La maîtrise est évidente.
Résultats corrects et visualisations cohérentes	[ 5 ]	Les calculs sont justes et le choix de chaque visualisation est parfaitement adapté à la question posée.
Structuration du notebook / workflow clair et reproductible	[ 5 ]	Le workflow est exemplaire et suit une démarche professionnelle. On peut relancer le notebook sans problème.
Pertinence de l'analyse exploratoire (EDA) et capacité à tirer des insights	[ 5 ]	L'EDA est le point fort du projet. Elle est très complète, chaque graphique est pertinent et accompagné d'une interprétation juste

Critère	Note /5	Commentaire
Capacité à répondre aux questions posées via des analyses/statistiques simples (NumPy/SciPy)	[5]	Très bonne utilisation des statistiques descriptives et inférentielles (test ANOVA) pour valider des hypothèses. Excellent.

---

#### ❏ **Commentaire final (optionnel)**

→ Excellent travail. Ce projet se distingue par sa rigueur méthodologique, la profondeur de son analyse exploratoire et la volonté d'aller plus loin avec des analyses statistiques et du clustering. Vous avez non seulement répondu à la problématique, mais vous avez aussi su raconter une histoire avec les données, en identifiant des profils et des facteurs de risque concrets. C'est un travail très abouti qui démontre une grande maturité analytique. Félicitations à toute l'équipe.